Perbandingan Convolutional Neural Network dan Convolutional Recurrent Neural Network sebagai Model Multiclass Keyword Spotting pada Edge Device = Convolutional Neural Network and Convolutional Recurrent Neural Network Comparison as Multiclass Keyword Spotting Model on Edge Device

Gita Ayu Salsabila, author

Abstrak

Selama masa pandemi COVID-19, antarmuka suara menggunakan KWS (keyword spotting) semakin sering digunakan pada berbagai sistem elektronik karena minimnya kontak fisik yang diperlukan antarmuka ini. Salah satu sistem yang dapat menggunakan KWS adalah sistem navigasi lift, di mana KWS pada sistem tersebut akan mengenali kata kunci terkait lantai yang ingin dituju pengguna. Dalam penelitian ini, model KWS untuk sistem navigasi lift dibuat menggunakan CNN (Convolutional Neural Network) dan CRNN (Convolutional Recurrent Neural Network) untuk mengenali enam kata kunci spesifik. Selama proses pembuatannya, berbagai hyperparameter CRNN terkait implementasi GRU, batch normalization, dropout layer, optimizer, kernel size, dan batch size diuji pengaruh variasinya terhadap performa CRNN. Dari pengujian tersebut, ditemukan bahwa CRNN menunjukkan performa paling baik ketika GRU yang digunakan bersifat bidirectional dengan dua layer dan 64 hidden unit, kernel size sebesar 3x3, optimizer Adams, batch size sebesar 163, serta penerapan batch normalization layer sebelum dropout layer. Model CRNN yang diperoleh dari kombinasi hyperparameter terbaik kemudian dibandingkan dengan model CNN untuk dievaluasi performa klasifikasinya saat dijalankan pada Raspberry Pi 4B. Berdasarkan hasil akurasi, persentase penggunaan RAM, dan latensi, model CNN menunjukkan performa yang lebih baik daripada CRNN.

.....During the COVID-19 pandemic, voice interfaces using KWS (keyword spotting) are increasingly being used in various electronic systems due to the lack of physical contact required for this interface. One system that can use KWS is an elevator navigation system, where the KWS on the system will recognize keywords related to the floor the user wants to go to. In this study, the KWS model for the elevator navigation system was created using CNN (Convolutional Neural Network) and CRNN (Convolutional Recurrent Neural Network) to identify six specific keywords. During the manufacturing process, various CRNN hyperparameters related to GRU implementation, batch normalization, dropout layer, optimizer, kernel size, and batch size were tested for the effect of their variations on CRNN performance. From these tests, it was found that CRNN showed the best performance when the GRU used bidirectional with two layers and 64 hidden units, kernel size of 3x3, Adams optimizer, batch size of 163, and batch normalization layer applied before dropout layer. The CRNN model obtained from the best combination of hyperparameters is then compared with the CNN model to evaluate its classification performance when run on the Raspberry Pi 4B. Based on the results of accuracy, percentage of RAM usage, and latency, CNN model shows better performance than CRNN.